Reproducible research with R

June 10, 2013

Part I

1. **Introduction**
   - Reproducibility
   - Tools needed
2. **Version control**
   - Interest of version control
   - Multiple users, single user
   - Version control depository
3. **Version control with R**
   - Setting up the project
   - example

Part II

4. **Introduction**
   - Report generation
   - Tools

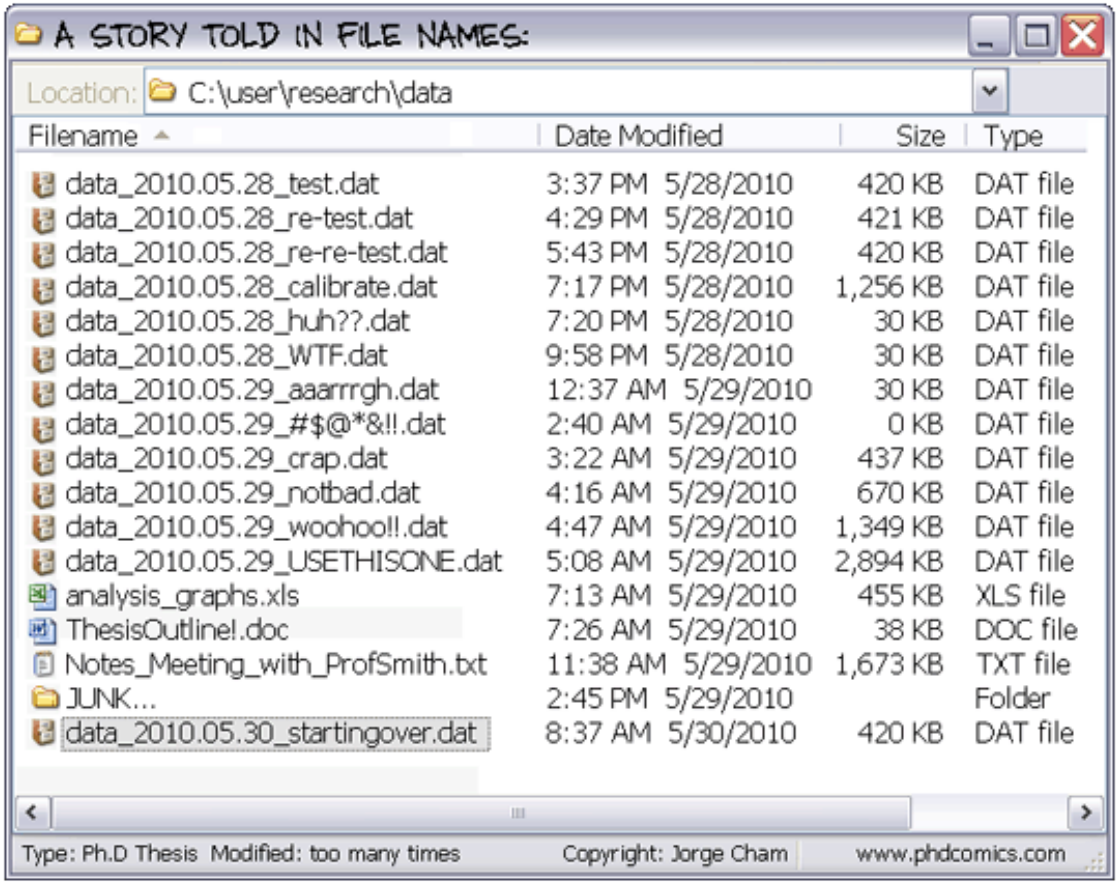
5. **Markdown - HTML output**

6. **LateX - PDF output**

Part I

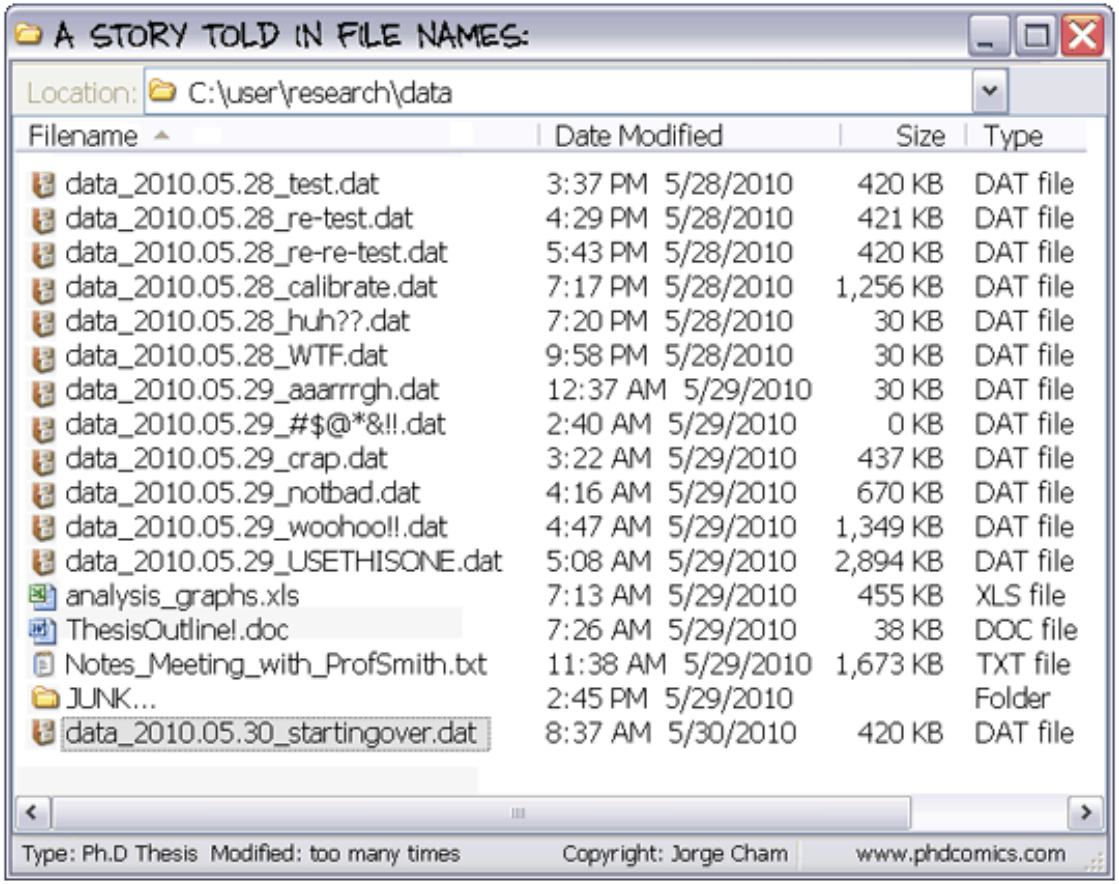
General concepts
Version control

A common problem



Not only with data, but also with scripts

A common problem



Not only with data, but also with scripts

Reproducibility ?

- Research should be reproducible by others.
- This refers to the experiments generating the data, but also to the analysis of the data.
- The first researcher who will need to reproduce the results is you.

Reproducibility ?

- Research should be reproducible by others.
- This refers to the experiments generating the data, but also to the analysis of the data.
- The first researcher who will need to reproduce the results is you.

Reproducibility ?

- Research should be reproducible by others.
- This refers to the experiments generating the data, but also to the analysis of the data.
- The first researcher who will need to reproduce the results is you.

Reproducibility of analysis

- Lab books make lab work traceable. Analysis should also be traceable.
- Analysis steps must be logged, and reverting to any previous step must be possible.
- This ensures that we always exactly know how a result was generated.

Reproducibility of analysis

- Lab books make lab work traceable. Analysis should also be traceable.
- Analysis steps must be logged, and reverting to any previous step must be possible.
- This ensures that we always exactly know how a result was generated.

Reproducibility of analysis

- Lab books make lab work traceable. Analysis should also be traceable.
- Analysis steps must be logged, and reverting to any previous step must be possible.
- This ensures that we always exactly know how a result was generated.

Version control

- Version control is a tool to keep track of file changes.
- However, version control softwares offer more than simply recording successive versions of a file.
- Version controlled projects can be splitted, merged and shared with collaborators.

Version control

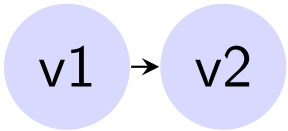
- Version control is a tool to keep track of file changes.
- However, version control softwares offer more than simply recording successive versions of a file.
- Version controlled projects can be splitted, merged and shared with collaborators.

Version control

- Version control is a tool to keep track of file changes.
- However, version control softwares offer more than simply recording successive versions of a file.
- Version controlled projects can be splitted, merged and shared with collaborators.

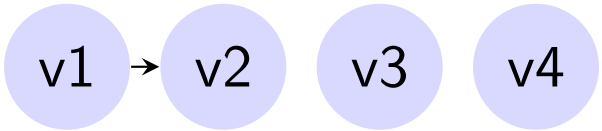
Version control flow

Version control flow

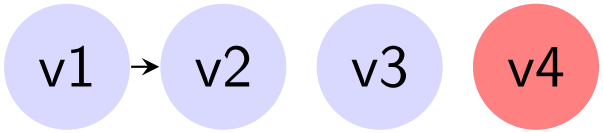


Version control flow

Version control flow

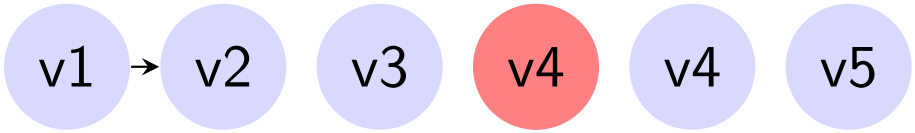


Version control flow

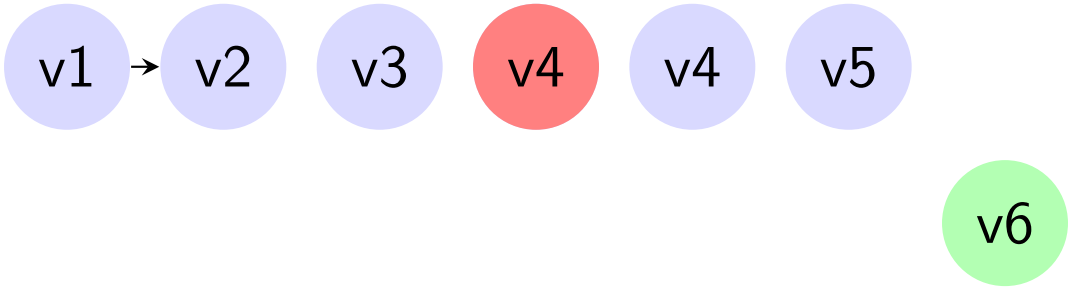


Version control flow

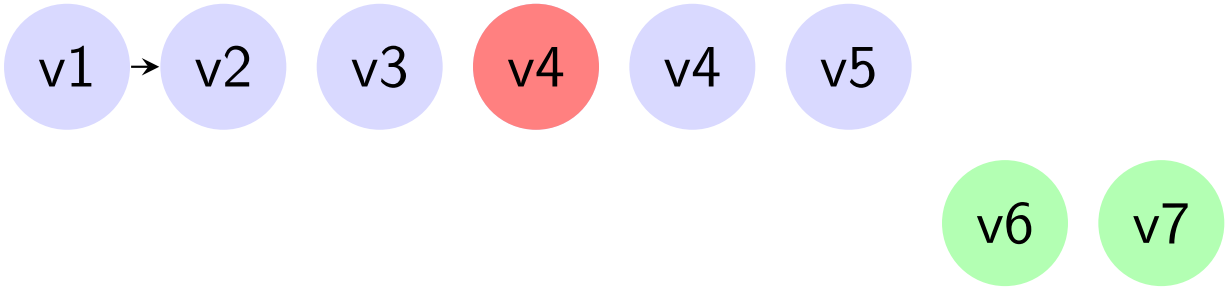
Version control flow



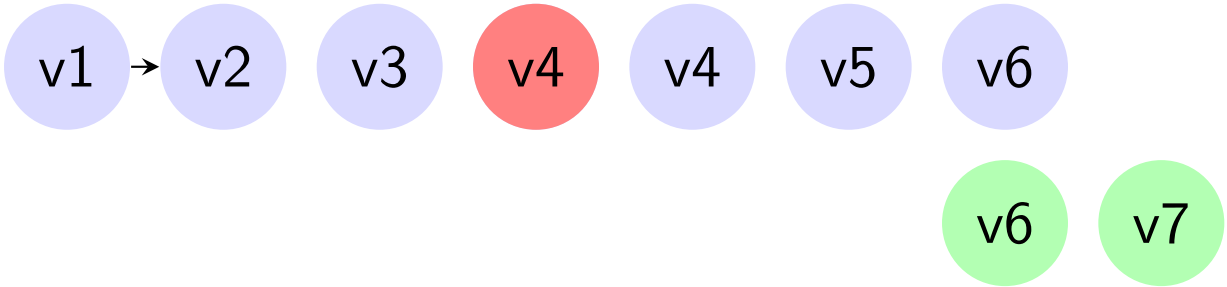
Version control flow



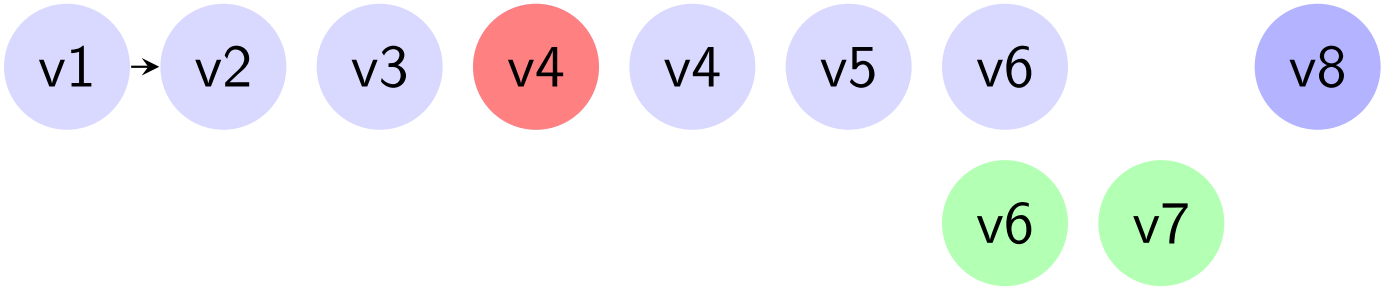
Version control flow



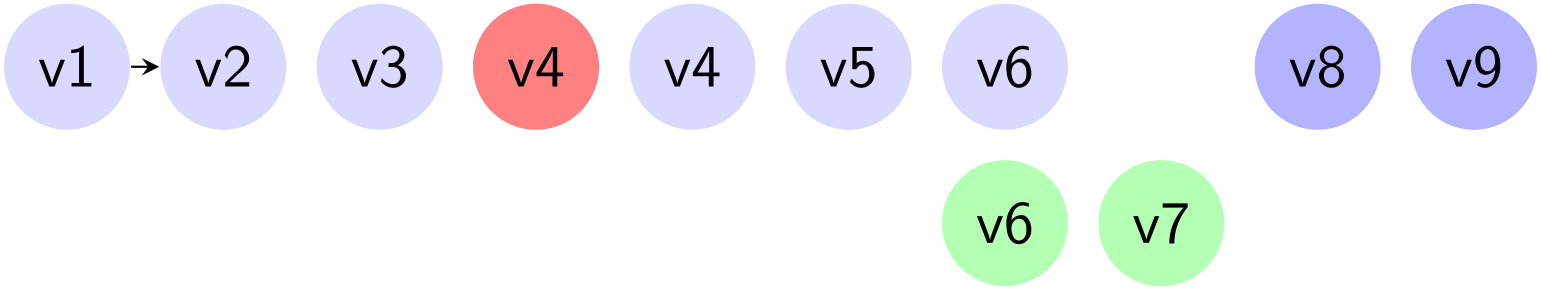
Version control flow



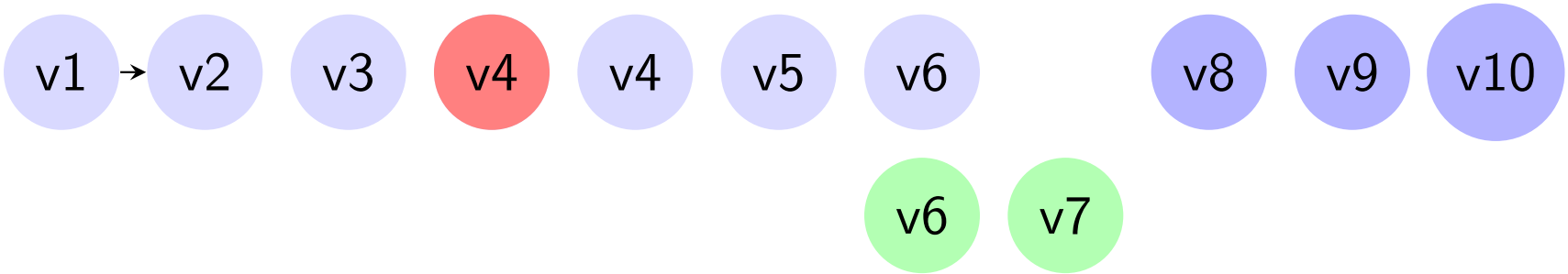
Version control flow



Version control flow



Version control flow



Example of project structure

blabla

Keeping track

What reproducibiity means

useful links and references

data : from http://ocean.ices.dk/HydChem/HydChem.aspx?plot=yes

Part II

Report generation with R
Figures with make

useful links and references